Cerebral Magnetic Resonance Image Segmentation Using Data Fusion

Jagath C. Rajapakse, Charles DeCarli, Alan McLaughlin, Jay N. Giedd, Amy L. Krain, Susan D. Hamburger, and Judith L. Rapoport

Objective: A semiautomated method is described for segmenting dual echo MR head scans into gray and white matter and CSF. The method is applied to brain scans of 80 healthy children and adolescents.

Materials and Methods: A probabilistic data fusion equation was used to combine simultaneously acquired T2-weighted and proton density head scans for tissue segmentation. The fusion equation optimizes the probability of a voxel being a particular tissue type, given the corresponding probabilities from both images. The algorithm accounts for the intensity inhomogeneities present in the images by fusion of local regions of the images.

Results: The method was validated using a phantom (agarose gel with iron oxide particles) and hand-segmented images. Gray and white matter volumes for subjects aged 20–30 years were close to those previously published. White matter and CSF volume increased and gray matter volume decreased significantly across ages 4–18 years. White matter, gray matter, and CSF volumes were larger for males than for females. Males and females showed similar change of gray and white matter volumes with age.

Conclusion: This simple, reliable, and valid method can be employed in clinical research for quantification of gray and white matter and CSF volumes in MR head scans. Increase in white matter volume may reflect ongoing axonal growth and myelination, and gray matter reductions may reflect synaptic pruning or cell death in the age span of 4–18 years.

Index Terms: Brain—Magnetic resonance imaging—White matter—Gray matter—Cerebrospinal fluid—Magnetic resonance imaging, physics and instrumentation.

The boundaries of anatomically and functionally distinct components of the brain are often defined by the junctures of white matter, gray matter, and CSF. Since brain structure is determined by these boundaries, the ability to accurately identify and quantify different tissue types is critical to quantitative studies of brain morphology. Analysis of changes in tissue structures and characteristics may provide important clues to understanding normal and abnormal brain development.

From the Child Psychiatry Branch, National Institute of Mental Health (J. C. Rajapakse, J. N. Giedd, A. L. Krain, S. D. Hamburger, and J. L. Rapoport), Epilepsy Branch, Neurological Disorders and Strokes Institute (C. DeCarli), and Clinical Brain Disorders Branch, National Institute of Mental Health (A. McLaughlin), National Institutes of Health, Bethesda, MD, U.S.A. Address correspondence and reprint requests to Dr. J. C. Rajapakse at Child Psychiatry Branch, National Institute of Mental Health, Bldg. 10, Rm. 6N240, 10 Center Dr., MSC 1600, Bethesda, MD 20982-1600, U.S.A.

While important measurements of microscopic brain anatomy can be done only with postmortem studies (1), postmortem material has several disadvantages: CSF is not available, and once removed from the intracranial cavity and extracerebral fluid, the brain collapses on its weight, distorting in vivo morphometric properties. CT techniques have been applied to separate gray matter, white matter, and CSF, but with limited success due to poor resolution (2,3). In contrast, MRI offers high spatial resolution and excellent contrast between gray and white matter (4). MRI is particularly well suited for studies of brain development since it uses no ionizing radiation and permits repeated scans.

Image segmentation is the process of separating image pixels into homogeneous regions of similar characteristics. A variety of segmentation methods are available (5,6). Brain MR image segmentation into gray matter, white matter, and CSF may be

done on single channel (7,8) or multiple channel data (9–23). Single channel data usually involve T1-weighted images, which give a good separation of gray and white matter. Multiple channel methods utilize more than one MRI property of tissues to enhance discriminating power and may provide a more sensitive basis for tissue differentiation.

Multiple channel data consist of a combination of two (9–16) or more (17–23) sets of images of T1-weighted, T2-weighted, and proton density (PD) images. Different techniques have been applied to handle multichannel MR image segmentation: (a) Two images are derived from the input images using discriminant analysis (9,10) or image algebra (11,12), one image to optimize the separation of brain matter and CSF and the other to optimize the separation of gray matter and white matter. A simple thresholding is then applied to separate different tissue types. (b) A set of low level rules is defined to classify different tissue types (13). (c) Tissue type separation is carried out using 2D histograms or scatter plots (14). (d) In multispectral analysis, the input data are considered to be multidimensional vectors where components in the vectors represent the corresponding voxel data from the images; the input vectors are then classified using a supervised or unsupervised classification system (15–20). (e) Input is considered as a multidimensional variable and the input variable is then classified using statistical approaches such as maximum likelihood and maximum a priori estimation techniques (21,22). (f) Multiple channel images are combined or processed to obtain more images or features to increase the input feature space for discrimination (23) and then classified using a higher dimensional classifier.

In this report, we present a data fusion technique based on a probabilistic framework to segment MR image data available in two channels in the form of T2-weighted and PD images. Data fusion refers to the combination or integration of more than one data set measuring the same physical event to get better information than that from any single set of data. Data fusion techniques have been previously considered in robotic vision, automatic target recognition, and medical imaging (24,25). Here, we introduce a probabilistic data fusion equation to combine dual echo MR images to segment cerebral tissues into gray matter, white matter, and CSF. This method functions to optimize the probability of a voxel location being a particular tissue type given information from both images.

A second objective of this study was to examine the maturation of white and gray matter and CSF during brain development in childhood and adolescence. Although most of the major anatomical development of the brain is completed in the first postnatal year, the configuration of the adult human brain continues to change throughout life (26,27). Previous quantitative studies of gray and white matter in the human brain have concentrated mostly on older age groups (1,10,11) and on neonates and infants (27,28), although a few studies have included our age group (29,30). Hence, quantitative analysis of tissue types in the understudied age span of 4–18 years permits scrutiny of brain development during the period of sexual maturation and across the age period in which childhood neuropsychiatric disorders emerge.

Although gray/white/CSF segmentation methods have been introduced in earlier studies, they have been applied only to smaller data sets. Most of them have been tried only on a small number of slices because of their instability when applied to larger data sets. As cross-sectional brain developmental data should necessarily be large to account for the complexity and individual variability of the human brain, an accurate and a stable segmentation method is needed that can be applied to a large data set for gray and white matter and CSF separation. We were motivated to introduce the present method because the available methods (10,11) did not produce consistent segmentation over our data base.

Reliability and a phantom validation study for the method are presented. Because of the unavailability of postmortem studies and limited imaging data for children, our gray and white matter and CSF volumes determined from segmentation of 13 MR head scans from patients ranging in age from 20 to 30 years were compared with previously published results for this age group.

SUBJECTS AND HARDWARE

MRI scans of brain were performed on 37 healthy male (mean age 11.98 ± 3.19 years) and 43 healthy female (mean age 10.87 ± 3.76 years) subjects with ages ranging from 4 to 18 years (mean age $\pm$ SD $= 11.38 \pm 3.53$ years). Individuals with physical, neurological, or lifetime history of psychiatric disorders were excluded. [Details of subject recruitment and demographics are presented elsewhere (31).]

All subjects were scanned using a GE 1.5 T Signa scanner. A variable echo multiplanar sequence in the axial plane was obtained (TE = 17 and 83 ms, TR = 2,600 ms, acquisition parameters = 192 × 256, NEX = 2, FOV = 24 cm) for the head. Slice thickness was 5 mm with 2.3 mm skip between successive cross-sections. The output images consisted of 36 interleaved 256 × 256 slices of PD-weighted images (TE = 17 ms) and T2-weighted images (TE = 83 ms). Voxel data have a resolution of 16 bits, which we converted into 8 bits before further processing to speed up processing and to reduce memory requirements. For volumetric analysis, the images were transferred to a Sparc 10 workstation and separated into PD- and T2-weighted slices.

PREPROCESSING

Our method segments only brain pixels, and hence it is necessary to separate the brain pixels from nonbrain pixels. Since PD images give the best resolution between the brain and the cranial boundary, PD images were selected to shell the brain. "Shelling" here refers to removal of all nonbrain material such as bone, skull, and face muscles from the images, leaving only brain material. It also removes the extracerebral CSF. Shelling was done using the auto-trace facility available in the software ANALYZE (32), in which a seed point is selected and then the user interactively manipulates a threshold slider to change the range of the threshold at the selected seed pixel. A trace is automatically drawn at the point where the threshold transition takes place, and the region inside or outside the trace may then be deleted from the image. Multiple seed points may also be specified. One or more seeds are planted on the brain matter, the threshold is adjusted for each slice, and nonbrain matter is deleted. The shelled PD images then serve to mask nonbrain tissues from T2 images. Shelled brain images include both cerebellum and brain stem.

FUSION SEGMENTATION

Theory

The input data can be represented as a set $\{P,T\}$ where P and T represent mappings of the PD image and the T2-weighted image, respectively. Segmentation is done slice by slice, one slice at a time. We will assume that the images are stacks of 2D images with rectangular pixel lattices of size $n \times m$, where n and m are integers. Then, if x is a pixel location in a given slice, $x \epsilon \{(0,0),(0,1), \ldots ,(n,m)\}$. We will denote the image intensity of the PD image at pixel location x as $P(x)$ and that of the T2-weighted image as $T(x)$.

We consider that the brain matter consists mainly of three tissues: gray matter, white matter, and CSF. Although most of the brain voxels are pure voxels, i.e., voxels containing only one type of tissue, the voxels at the tissue boundaries may contain more than one tissue type. These mixed pixels create a statistical error causing partial volume effects, which will be discussed later. When a particular voxel is occupied by more than one tissue type, it is assumed that the voxel will be occupied completely by the most probable tissue type. Consequently, we can think of image pixels in a given image slice consisting of three mutually exclusive and exhaustive sets: white matter (W), gray matter (G), and CSF (F). W, G, and F are disjoint sets of pixels in the image.

Statistically, the feature that distinguishes one tissue from the other is the probability distribution of the intensity that each tissue type produces over the image. Since the pixel intensities are influenced by random noise and correlated with the physical relaxation processes that are responsible for creating the intensity values, we assume Gaussian distributions for the probability distribution of intensities in each tissue (8). DeCarli et al. (33) used a skewed Gaussian distribution for their single channel images to account for the partial volume effect. However, it will be shown later that the present algorithm partially accounts for the volume averaging effect.

The statistical parameters that are associated with tissues can be characterized by means and standard deviations of the assumed Gaussian distributions. We denote the mean and the standard deviation of the tissue class C in image I at pixel location x by $\mu_{I,C}(x)$ and $\sigma_{I,C}(x)$ respectively, where $C \epsilon \{W,G,F\}$ and $I \epsilon \{P,T\}$. The probability that the pixel at location x in image I belongs to tissue class C is denoted by $p_{I,C}(x|I)$. The intensity in image I at location x is denoted by $I(x)$. With our notation and the Gaussian distribution assumption, the probability distribution of tissue types in the two images can be written as

$$p_{I,C}(x|I) = \frac{1}{\sqrt{2\pi\sigma^2_{I,C}(x)}} \exp^{-\frac{1}{2}\left[\frac{I(x)-\mu_{I,C}(x)}{\sigma_{I,C}(x)}\right]^2} \quad (1)$$

where $I \epsilon \{P,T\}$ and $C \epsilon \{W,G,F\}$.

Equation 1 defines the probabilities of a particular pixel being white or gray matter or CSF, given image intensity and distribution parameters. Some means is needed for combining these probabilities to obtain the probabilities for each tissue type given both sets of image data. The fusion equation (proved in Appendix A) provides a means of combining data probabilistically when independent information is available from more than one set of data. We will approximately use the fusion equation to combine probabilities implied by both images. If $p_C(x|P,T)$ is the probability of a pixel at location x being tissue class C given both PD image and T2-weighted image, then the fusion equation can be written as

$$p_C(x|P,T) = k_1 \cdot p_{P,C}(x|P) \cdot p_{T,C}(x|T) \quad (2)$$

where k_1 is a real positive constant and $C \epsilon \{W,G,F\}$.

The fused probability given by Eq. 2 refers to the probability of the pixel at location x being class C jointly implied by intensities of both the PD image and the T2 image. This is different from the probability calculated as $p_C(x|y)$ with y being a multidimensional vector, $y = [P(x),T(x)]$, as considered in the multispectral analysis or multidimenstional feature approach (15–20,23,34). The fused probability

$p_C(x|P,T)$ calculated by multiplying the individual probabilities given by each sensor gives a probability estimate that is better than $p_C(x|T)$ and $p_C(x|P)$ for classification. This value is theoretically optimum when the two images are fully independent.

Tissue type is assigned to a particular pixel location according to maximum likelihood criteria (35). That is, the tissue type is the one whose probability is the highest given the image data from both the images. Then, pixel location x is considered to be white matter if

$$p_W(x|P,T) > p_G(x|P,T) \text{ and}$$
$$p_W(x|P,T) > p_F(x|P,T) \qquad (3)$$

and it is considered to be gray matter if

$$p_G(x|P,T) > p_W(x|P,T) \text{ and}$$
$$p_G(x|P,T) > p_F(x|P,T) \qquad (4)$$

Otherwise it is considered as CSF.

Since the assignment of tissue type is made using the information from both images, greater accuracy is achieved than with information from only one source. The fusion equation always gives a better probability value for a pixel to be particular tissue type than given by the individual images. The fusion equation could be easily applied to combine more than two input images when they can be considered independent.

Application to Dual Echo MR Images

PD images and T2-weighted images in a double echo sequence are spatially registered since they are acquired simultaneously using the same pulse sequence and therefore are suitable for probabilistic data fusion assuming two sensor channels. Under the present imaging conditions, PD and T2-weighted images are not purely independent. The lack of complete independence, however, does not invalidate the fusion equation, which always gives a better probability than that obtained by the individual images, as we will show later. Any dependencies between the images act as redundant information and lower the efficiency of the fusion process.

At the present time, we are interested primarily in the segmentation of gray matter and white matter in normal brains (i.e., brains without the presence of any abnormalities such as tumors). In this regard, we assume a three class model, where CSFs in the two images are orthogonal. We recognize that other brain tissue types exist, however. Other brain tissue types include blood vessels and hyperintense white matter signals. Due to the low signal values on PD and T2 image data, vascular structures in the brain will be classified as white matter. The percent error due to this misclassification is, however, very small. In the present version of this method, hyperintense white matter signals are misclassified as gray matter. The presence of hyperintense white matter lesions did not significantly affect our results, as clinical review of each MR image revealed a low prevalence of white matter hyperintense lesions in our subject population.

Because of the inhomogeneities in the magnetic field and RF coil sensitivity of the imaging scanner, regional intensity variations are often present in MR image slices. This has been a major obstacle for segmentation of the tissues based on intensities of the pixels. Our images did not show severe inhomogeneity artifacts, and the subtle inhomogeneities present seemed smoothly varying. The parameters of the distributions were locally evaluated by considering a local neighborhood at the pixel as the neighborhood pixels are equally affected by the inhomogeneities present in the scanner. The algorithm does not depend on a global threshold or any parameters computed over the whole image. Instead, the parameters for the Gaussian distributions are calculated by considering a local neighborhood window at each pixel. Since the distributions are determined at each pixel separately, smooth spatial variations of intensities do not affect the segmentation. Further, slice-by-slice segmentation alleviates problems with intensity drifts in the axial direction. As expected, the segmented images did not show the effect of magnetic field or RF inhomogeneities.

Implementation

The fusion algorithm is implemented in an iterative manner. In each iteration, the decisions for tissue classes are made at every pixel location by comparing the probability distributions of tissue types as in Eqs. 3 and 4. The probabilities are calculated using the fusion Eq. 2 and the individual probabilities are computed using Eq. 1. The parameters of the distributions are initialized at the beginning of the iterations, and after each decision about the tissue class is made, the parameters of distributions are updated. At each pixel location, parameters for the distributions are calculated considering a local neighborhood D, which consists of pixel locations with a window of size 25×25.

Initially, the distribution parameters for the tissues are determined by the parameters of the two input images. If $\mu_P(x)$ and $\sigma_P(x)$ are the mean and the standard deviation of the intensities at pixel location x in PD image, then the parameters of the PD images are initialized as follows:

$$\mu_{P,G}(x) = \mu_P(x) + k_2 \cdot \sigma_P(x) \qquad (5)$$

$$\mu_{P,W}(x) = \mu_P(x) - k_2 \cdot \sigma_P(x) \qquad (6)$$

$$\mu_{P,F}(x) = k_2 \cdot \sigma_P(x) \qquad (7)$$

$$\sigma_{P,G}(x) = \sigma_{P,W}(x) = \sigma_{P,F}(x) = k_3 \cdot \sigma_P(x) \qquad (8)$$

Similarly, if $\mu_T(x)$ and $\sigma_T(x)$ are the mean and the standard deviation of the intensities of the T2 image, the parameters of the T2 image are initialized as follows:

$$\mu_{T,G}(x) = \mu_T(x) + k_2 \cdot \sigma_T(x) \tag{9}$$

$$\mu_{T,W}(x) = \mu_T(x) - k_2 \cdot \sigma_T(x) \tag{10}$$

$$\mu_{T,F}(x) = I_{max} - k_2 \cdot \sigma_T(x) \tag{11}$$

$$\sigma_{T,G}(x) = \sigma_{T,W}(x) = \sigma_{T,F}(x) = k_3 \cdot \sigma_T(x) \tag{12}$$

where k_2 and k_3 are real positive constants and I_{max} is the maximum intensity of the image. $k_2 = k_3 = 0.5$ worked well with our images. k_2 and k_3 determine the initial separation and shapes of the distributions. It should be noted here that the gray matter is brighter than white matter in both images and CSF is the brightest tissue type in the T2 image and the darkest in the PD image. Representational slices of the PD and T2-weighted images are shown in Fig. 1.

When a decision is made at a pixel location to be of particular tissue class, the parameters of the distribution for that tissue class are updated. When $n_C(x + v)$ is the number of pixels already classified as tissue class C, at the pixel location $x + v$ where $v \in D$, and a decision at pixel location x is taken to be of tissue type C, then the parameters of the distributions for the PD image and the T2 image for tissue class C are updated within a window D, as follows:

$$\mu_{P,C}(x + v) = \frac{n_C(x + v) \cdot \mu_{P,C}(x + v) + P(x)}{n_C(x + v) + 1} \tag{13}$$

$$\mu_{T,C}(x + v) = \frac{n_C(x + v) \cdot \mu_{T,C}(x + v) + T(x)}{n_C(x + v) + 1} \tag{14}$$

$$\sigma^2_{P,C}(x + v) =$$

FIG. 1. Representative MR brain image slices: proton density image slice **(a)**; T2-weighted image slice **(b)**.

$$\frac{n_C(x + v) \cdot \sigma^2_{P,C}(x + v) + [P(x) - \mu_{P,C}(x + v)]^2}{n_C(x + v) + 1} \tag{15}$$

$$\sigma^2_{T,C}(x + v) =$$

$$\frac{n_C(x + v) \cdot \sigma^2_{T,C}(x + v) + [T(x) - \mu_{T,C}(x + v)]^2}{n_C(x + v) + 1} \tag{16}$$

$$n_C(x + v) = n_C(x + v) + 1 \tag{17}$$

where $C \in \{W,G,F\}$ and $v \in D$. Variations of this parameter updating scheme have previously been used by a number of classification algorithms. A global version of a similar parameter updating scheme was used in unsupervised K-means algorithm (34). In the global version, one set of distributions is considered over the whole image, and the parameters of these distributions are updated at each point whenever a decision is made. The global scheme runs much faster than the local updating of the variables, but is sensitive to local changes of intensity over the image.

The steps of the implementation of the segmentation procedure are shown in algorithmic form herein for implementation in a high level language.

For every slice in the image
   Change of error = HIGH
   Find values μ_P and σ_P from PD image and μ_T and σ_T from T2 image
   Initialize $\mu_{I,C}$ and $\sigma_{I,C}$ for $C \in \{G,W,F\}$ and $I \in \{P,T\}$ using Eqs. 5–12 at each point
   While (change of error $\geq \epsilon$)
      For each pixel location x in the image
         Compute $p_{I,C}(x|I)$ for $I \in \{P,T\}$, and $C \in \{W,G,F\}$ using Eq. 1
         Using the fusion equations in Eq. 2, compute $p_W(x|P,T)$, $p_G(x|P,T)$, and $p_F(x|P,T)$
         If $[p_W(x|P,T) > p_G(x|P,T)]$ and $[p_W(x|P,T) > p_F(x|P,T)]$
            Assign pixel x white matter label (W)
            Update $\mu_{P,W}$, $\mu_{T,W}$, $\sigma_{P,W}$, $\sigma_{T,W}$, and n_W within a window D, as in Eqs. 13–17
         Else if $[p_G(x|P,T) > p_W(x|P,T)]$ and $[p_G(x|P,T) > p_F(x|P,T)]$
            Assign pixel x gray matter label (G)
            Update $\mu_{P,G}$, $\mu_{T,G}$, $\sigma_{P,G}$, $\sigma_{T,G}$, and n_G within a window D, as in Eqs. 13–17
         Else
            Assign pixel x CSF label (F)
            Update $\mu_{P,F}$, $\mu_{T,F}$, $\sigma_{P,F}$, $\sigma_{T,F}$, and n_F within a window D, as in Eqs. 13–17
      Continue for all pixels in the image
      Error = mean square value of the difference between tissue means and pixel intensities
   Continue until the change of error is small enough
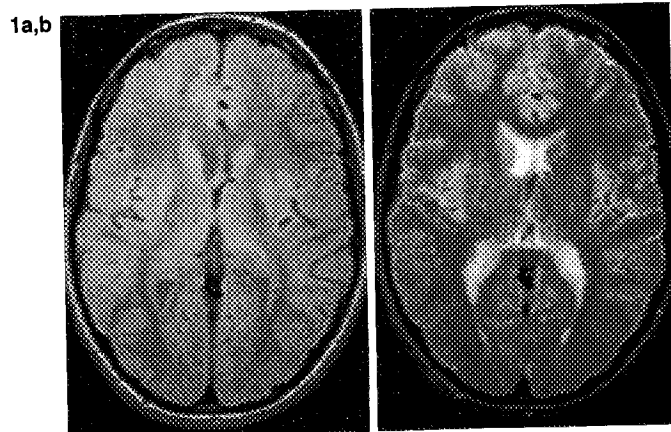Continue for all the slices in the image.

The error is computed as the mean square error between the mean tissue intensity and the intensity of the pixel. The classification of pixels over an image slice continues until the error stabilizes and the change of error drops below a certain small value. For most of the images, after three or four iterations over the images, the error became stabilized.

At the end of the segmentation, gray and white matter and CSF images are stored separately and the volumes are computed. It should be noted that brain volume here includes both cerebellum and the brain stem. The brain stem extends down to the foramen magnum.

RESULTS

Reproducibility

The reproducibility of our method depends to a large degree on the shelling process of the head scan. Shelling was done by manually adjusting the intensity threshold until there was enough resolution between the brain and the intracranial cavity. Since the rest of the method is automated, the reliability of the segmentation method is affected only by the brain shelling process. To assess the intraoperator reliability, the same operator segmented 10 scans of head twice. The interclass correlation coefficient values for the intrarater reliabilities were gray matter $= 0.95$, white matter $= 0.94$, and ventricular CSF $= 0.89$. To obtain the interrater reliability, the same 10 images were shelled by two different raters and then applied to the segmentation algorithm. Interclass correlation coefficient values for interrater reliabilities for gray matter, white matter, and CSF were 0.94, 0.92, and 0.88, respectively.

Validation Against Hand-Traced Images

To validate our results and to examine data from our age range, another segmentation was made with manual tracing of gray matter and white matter guided by one of the authors with extensive experience in brain image analysis. For this purpose, three slices were separated from a representative brain, and gray and white matter and CSF regions were manually outlined on each slice. The hand-segmented tissue volumes and the volumes obtained by the fusion segmentation algorithm are shown in Table 1 with their differences.

Phantom Validation

Validation requires a composite phantom, with the relative amplitude of the MRI signals from different regions approximating that for gray matter, white matter, and CSF in both PD and T2 images. We made up agarose gels (9,36) (1% wt/vol) with varying concentrations of dextran-coated superparamagnetic iron oxide particles (AMI-25 particles; Advanced Magnetics) having an average particle diameter of 500 Å. These particles are soluble, do not aggregate, and are large enough so that they should not diffuse across gel boundaries. After experimenting with test phantoms, we found that 1% (wt/vol) agarose gel with total iron concentrations of 6.7, 33, and 67 μM were suitable for simulating CSF, gray matter, and white matter, respectively. Increasing iron concentration decreased T1 and T2 (37). The decrease in T2 caused a decrease in the intensity in the T2 image. The decrease in T1 caused an increase in the intensity in the PD image, because in the absence of iron particles the signal was partially saturated.

The phantom consisted of a cylindrical plastic bottle with layers of 1% agarose containing different iron oxide concentrations. The layers were formed by pouring molten agarose gel containing a specific iron oxide concentration into the bottle, allowing it to gel, and then pouring molten agarose with a different iron oxide concentration on top. We chose this simple configuration because our segmentation method is "pixel based" and is not sensitive to structural details.

Dual echo images of the phantom were acquired in a plane parallel to the long axis, under the same conditions used for human subjects. One pair of slices through the same region of the PD and T2 images is shown in Fig. 2. Images were segmented by the fusion segmentation algorithm, and the volumes of different layers were computed. Actual and

TABLE 1. *Validation of gray matter and white matter obtained by fusion segmentation method versus hand tracing*

Slice level	Gray matter			White matter			CSF		
	Manual	Fusion	Difference (%)	Manual	Fusion	Difference (%)	Manual	Fusion	Difference (%)
Midventricular	71.00	69.70	1.8	62.04	63.47	−2.2	2.70	2.57	4.8
Basal ganglia	56.72	55.39	2.3	60.92	62.04	−1.8	2.19	2.41	−9.8
Lower cortex	48.09	50.01	3.9	49.83	47.55	4.6	1.16	1.32	13.4

The volumes are in voxels. Percentage difference = ratio of (hand-traced volume − volume obtained by fusion algorithm) to hand-traced volume.
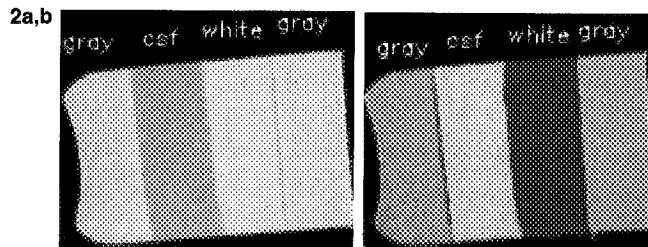
2a,b

FIG. 2. Phantom images: (a) Proton density image slice and (b) T2-weighted image slice taken parallel to the long axis of the phantom.

measured volumes of different gel layers are shown in Table 2.

Pediatric MR Data

The segmentation algorithm was implemented in the C/Unix environment and applied to the MR brain scans of 80 normal children (43 females and 37 males) aged 4–18 years. Each head image was shelled slice by slice, which took ~10 min per brain on our Sparc 10 workstation. After head scans were shelled, both images were applied to the segmentation program. The output gray matter, white matter, and CSF volumes were stored and measured. Samples of the segmented images of gray matter, white matter, and CSF of three slices are shown in Fig. 3.

Probability maps over image slices taken at a particular level of the brain, before and after fusion, are shown in Fig. 4. In probability maps, intensity of a pixel is proportional to the probability of that pixel being the corresponding tissue type. Maps of PD and T2 images were obtained before the fusion took place, assuming local Gaussian distributions as considered in Eq. 1. As seen, the T2 image better represents white matter, whereas the PD image better represents gray matter. The fused image probabilities optimize for classification of both gray matter and white matter.

Scatterplots and regression lines for CSF, gray matter, and white matter volumes in milliliters versus age are shown in Fig. 5a, b, and c, respectively, for 80 subjects. As seen, there is significant age-

TABLE 2. *Actual and measured volumes of different regions of phantom*

Sample region	Actual volume (ml)	Measured volume (ml)	Difference (%)
White	200.0	198.0	1.0
CSF	200.0	194.7	2.6
Gray	400.0	407.3	1.8

The different regions were composed of 1% agar gels with varying concentration of dextran-coated iron oxide particles. The total iron concentrations in the different regions were "white matter," 67 μM; "CSF," 6.7 μM; "gray matter," 33 μM.
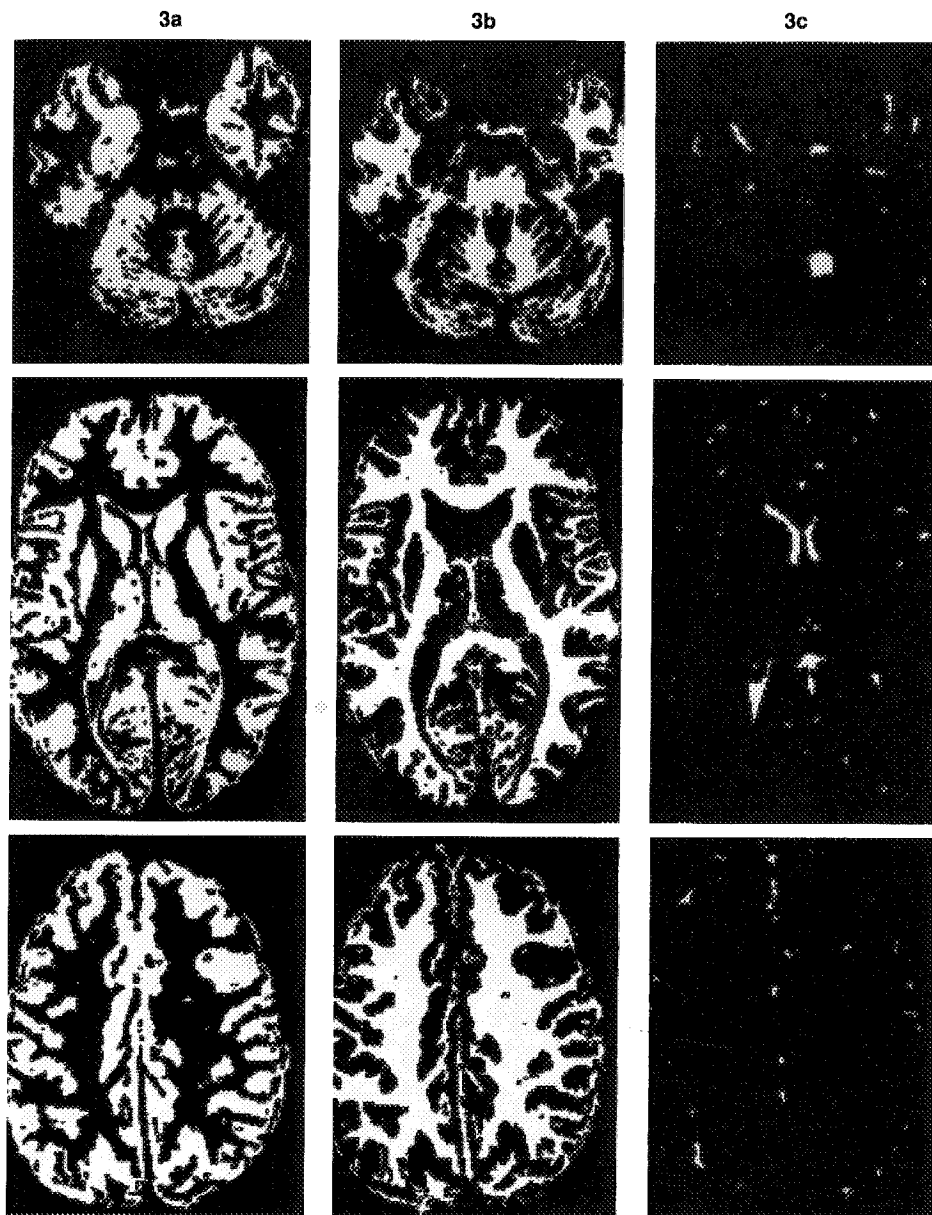
related increase in white matter volume (p < 0.001) and CSF volume (p < 0.001) for females. The decrease in gray matter volume was not significant. In males, the increase of CSF volume was significant (p = 0.001), but neither the white matter increase nor the gray matter decrease reached significance. The gray matter/white matter volume ratios were computed as shown in Fig. 5d. This ratio decreased significantly with age (for females p < 0.001 and for males p = 0.003).

To account for the effect of brain size on tissue volumes, the gray matter, white matter, and CSF volumes were normalized by dividing each volume by the total brain size. When this was done, the white matter/brain ratio significantly increased (p = 0.026 for males and p < 0.001 for females) and the gray matter/brain ratio significantly decreased (p < 0.001 for females and males) with age. The CSF/brain ratio increased significantly with age both in males and in females (p < 0.001). The regression coefficients were all linear, and the higher order regression coefficients were not significant.

The slopes of the regression lines for gray matter volume ($t = 3.29$, p = 0.004), white matter volume ($t = 3.93$, p = 0.001), and total brain volume ($t = 5.50$, p < 0.001) were significantly larger for males than females across all ages. There were no significant differences between the slopes of any of these regression lines calculated for males and females when corrected for the brain size.

The correlation coefficients and p values for these variables with age for each gender and for the total group are summarized in Table 3. All the variables we considered were highly and significantly correlated with age when ratios were considered with respect to the total brain volume.

DISCUSSION

A promising method is presented for segmenting white matter, gray matter, and CSF volumes in MR head scans using a probabilistic fusion equation. Although simultaneously obtained and thus automatically registered T2 and PD images were not fully independent, the use of the fusion equation resulted in accurate segmentation data. Hence, we believe that the fusion equation gives better estimates of the probabilities to differentiate tissue volumes combining information in dual echo images than those given by individual images. As Fig. 4 illustrates, a clear distinction among tissue types was seen when the probabilities of the tissue types were fused.

Since the segmentation is automated, the method is free from the subjective errors of operator thresholding prevalent in the previous approaches. Previous methods combined the two input images to create an image that was better or optimal for separation of gray matter and white matter (9–12), and

3a   3b   3c

FIG. 3. Segmented image slices taken at three different levels of the brain: gray matter **(a)**; white matter **(b)**; CSF **(c)**.

thresholding was performed to separate gray and white matter. In this report, we describe a somewhat different strategy using a probabilistic data fusion equation that computes a better probability estimate of a pixel by combining the probabilities implied by both input images. The parameters of the probability distributions are obtained by utilizing the full extent of local pixel intensities as opposed to simple or "idealized" distributions using a few sample points.

When only PD and T2 images are available, white and gray matter and CSF cannot be fully separated (19) using a 2D multispectral analysis. Multispectral analysis requires at least three uncorrelated images. Probabilistic fusion equation gives an optimal result when the input data sets are statistically indepen-

dent, while clustering in multidimensional space is optimum when the input data sets are orthogonal in the feature space. When registered images are acquired simultaneously using the sequence described here, it is difficult to obtain purely independent data. One may want to manipulate TR and TE values of the scanning parameters to obtain nearly independent and registered images to get the full advantage of the fusion equation.

There are unique advantages of this method of segmentation. Because the method is unsupervised, it does not need operator selection of training samples for classification. No parameters are set by the operator, and the final outcome is less sensitive to the preset parameters k_2 and k_3. The results are independent of k_1. Previous unsupervised techniques

4a   4b   4c

FIG. 4. Probability maps for tissue categories of individual (**a:** proton density; **b:** T_2-weighted) and (**c**) fused image slices. Intensity at a pixel location in a probability map represents the probability of that pixel being the corresponding tissue type. Row 1, white matter; row 2, gray matter; row 3, CSF.
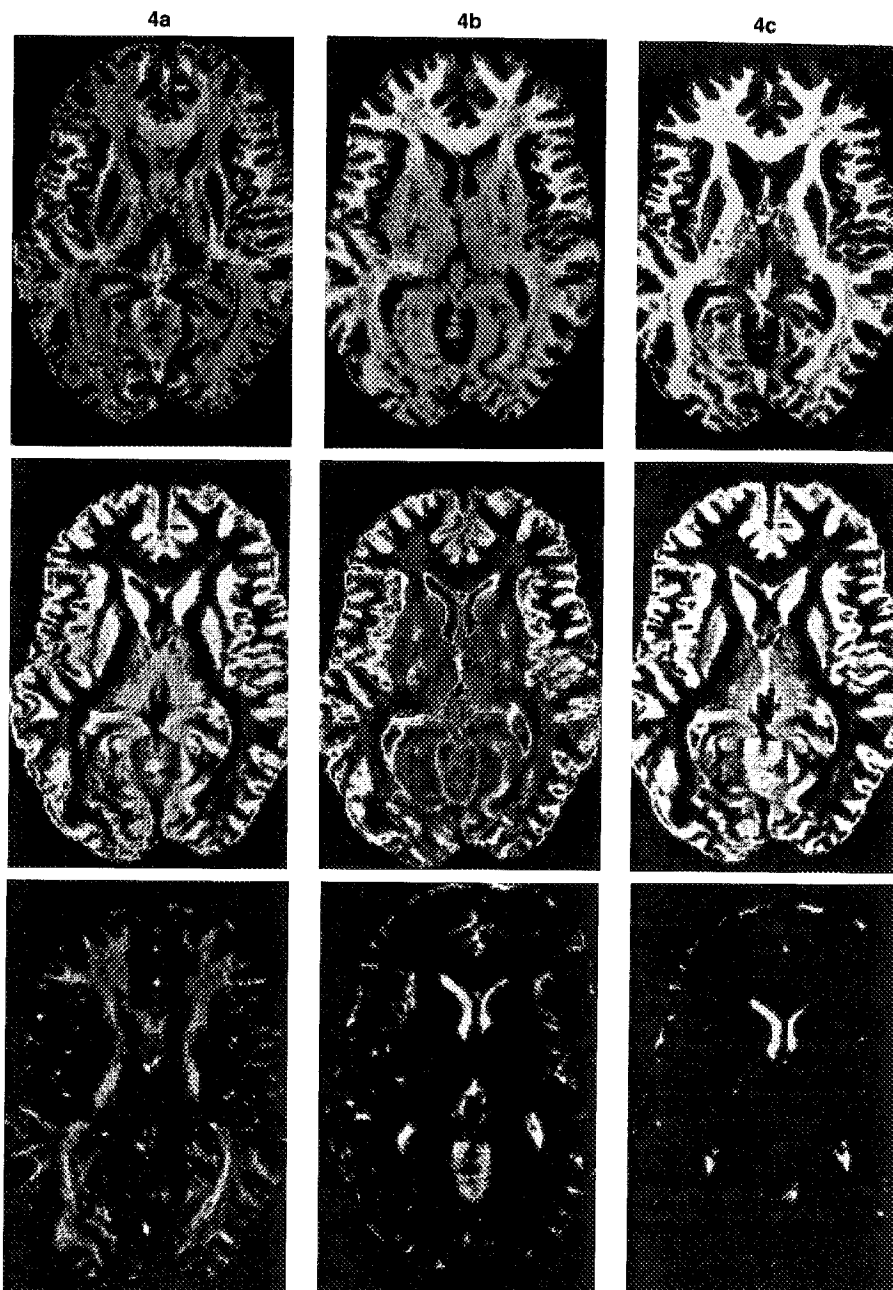
used for the same purpose have not yielded accurate segmentation and were limited to segmentation of cortical gray and white matter (29,30), as most previous methods classified some of the subcortical gray matter nuclei as white matter. The results of our segmentation were visually correct and all subcortical gray matter nuclei were classified as gray matter; this method could be easily extended if a third independent and registered T1 data set were available. Since it does not assume global parameters, the method accounts for the image intensity inhomogeneities. The parameters of the distributions are locally calculated by considering neighborhood pixels at each pixel location and not dependent on global distribution of intensity. Furthermore, the algorithm makes efficient use of computer processing time and does not require sophisticated mathematical modeling software.

The method's reproducibility was demonstrated by the acceptable intrarater and interrater reliabilities. The reliability of the method depends only on the brain shelling portion of the algorithm, which
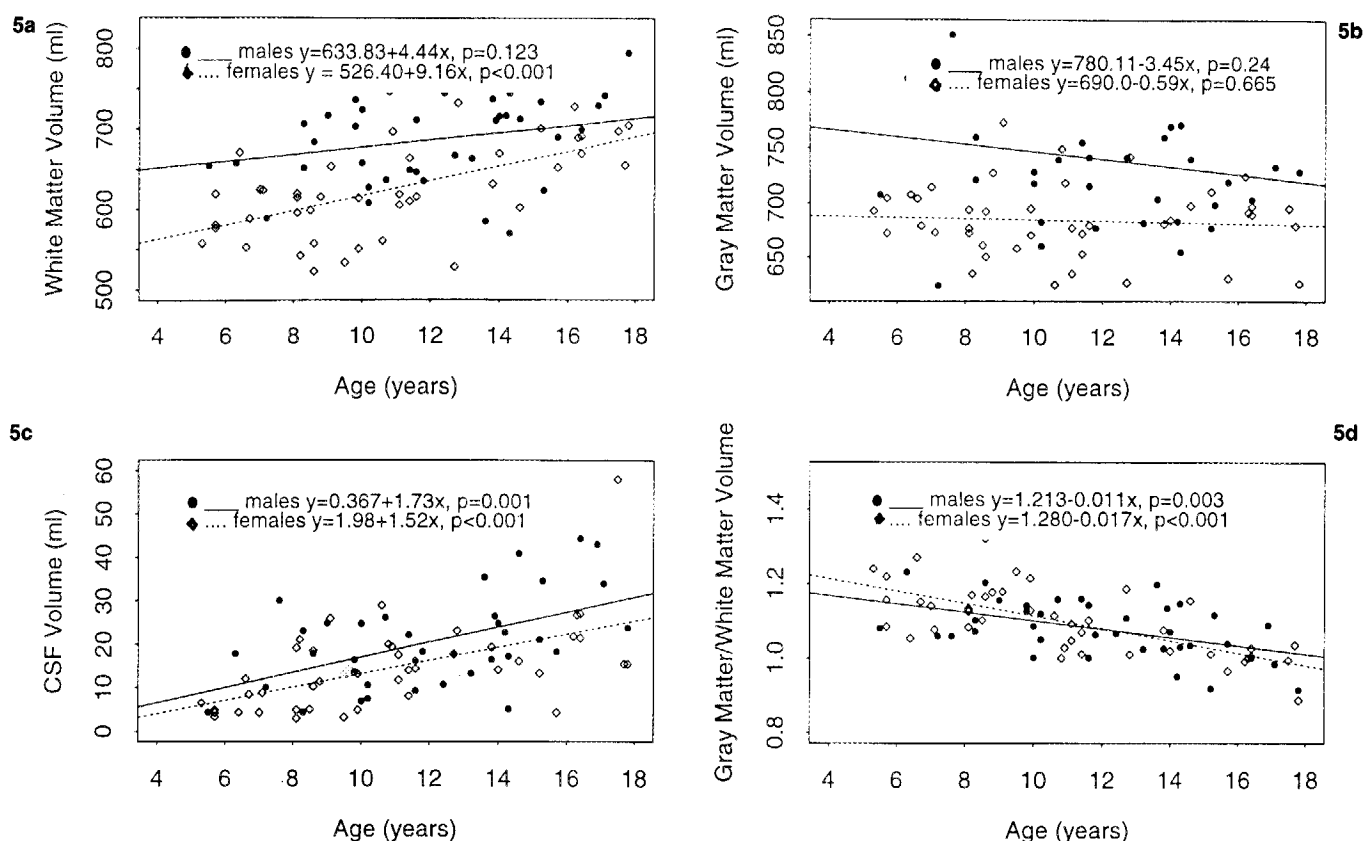
FIG. 5. a: Plot of white matter volume (ml) against age (yrs). **b:** Plot of gray matter volume (ml) against age (yrs). **c:** Plot of CSF volume (ml) against age (yrs). **d:** Plot of ratio of gray to white matter volume against age (yrs).

may further be improved with better brain shelling technique. Validation of the method was shown using both manual tracing technique and phantom studies. The fusion algorithm gave visually better segmentation than did the hand-segmented images. Validation with the phantom was good, as the errors for the tissues were <5% for all types of tissues.

There are no reports in our sample's age group for comparison with these segmentation data. Pfefferbaum et al. (29) analyzed only cortical gray and white matter and CSF. Hence, for comparison, we were limited to the 13 normal adults between age 20 and 30 years. Miller et al. (1) has used fixed brain sections with digital image analyzer to perform analogous postmortem studies across ages 20–98 years. For their subset, eight subjects aged 20–30 years, the gray and white matter ratio varied between 1.12 and 1.5 with a mean of 1.23 (obtained from the graph for cerebrum only). Lim and Pfefferbaum (11) also segmented subjects aged 20–30 years, obtaining a gray/white matter ratio of 1.12.

TABLE 3. *Correlation of tissue volumes for males, females, and total group with age*

	Males		Females		Total	
	r	*p*	*r*	*p*	*r*	*p*
Gray volume	−0.197	0.242	−0.068	0.665	−0.028	0.809
White volume	0.254	0.128	0.583	<0.001	0.465	<0.001
CSF volume	0.521	0.001	0.566	<0.001	0.559	<0.001
Gray/brain volume	−0.587	<0.001	−0.749	<0.001	−0.697	<0.001
White/brain volume	0.367	0.026	0.618	<0.001	0.524	<0.001
CSF/brain volume	0.529	<0.001	0.543	<0.001	0.550	<0.001
Gray/white volume	−0.477	0.003	−0.682	<0.001	−0.612	<0.001

r denotes the Pearson correlation coefficient.

Our value of 1.14 for gray/white matter ratio for this age range closely approximates these previously published data. Recently, Harris et al. (12) reported a similar segmentation study with manual thresholding. Their values for the CSF, white matter, and gray matter percentage for the age range 10–20 years were 5.83, 42.6, and 52.0% (obtained from regression lines), close to our values of 5.0, 42.9, and 52.1% for the same variables, respectively.

The volume averaging effect at the boundaries between gray and white matter and CSF is partially accounted for by the segmentation algorithm. CSF appears as the brightest tissue type in T2 images and the darkest in PD images. In both images, white matter appears darker than the gray matter. Therefore, partial volume effects in PD images appear at the CSF and gray matter boundaries and in T2 images at the CSF/white matter boundary. When the images are fused, the partial volume effects partially cancel out at the boundaries, because partial volume averaging is effective only in one image at any boundary (see Fig. 4). The pixels that contain multiple tissue types in one image usually give a lower probability compared with "pure" tissue pixels in the other image. So, the pure pixels dominate in the fusion equation driving the final classification. So, incorrect probabilities given by the pixels containing "mixed" tissues affect the segmentation less. Still, partial volume effects are seen at the brain–background boundary and caused a ringing effect on our segmented images.

A disadvantage of our present method is the use of thick slices (5 mm) for segmentation, which can enhance both partial averaging effects at the boundaries of brain and intracranial cavity and errors in computation of volumes. A phantom study has shown that the errors in volume computation and the partial volume effects increase with the thickness of the slices (38). We used the same slice thickness for all the subjects irrespective of their brain sizes as there was no significant brain size change with age in our age span (31). Our method was slowed by the semiautomated approach for brain shelling and calculation of parameters at every voxel location. With use of 18 slices, the shelling and segmentation process takes only ~25 min per brain on our Sparc 10 workstation; more numerous thinner slices would, of course, require more time.

In this study, the total brain volume included cerebellum and brain stem. With the segmentation of MR brain scans of 80 normal children and adolescents, both the white matter and the total brain volume increased (significantly for females), and the ratio of gray matter to white matter decreased. The clear increase in cerebellar volume in this age span (31), which may account for the increased brain volume found here, is consistent with earlier findings of increasing brain weight across our age span (39,40).

At birth, both myelinated and unmyelinated neurons are present in the human brain, and myelination of neurons continues as the brain develops. Axon myelination by oligodendrocytes results in an increase in myelinated white matter (41). Our findings of decreased gray/white matter ratio through adolescence are consistent with these earlier findings and the recent MR study (30). Visual examination of segmented images at various age levels showed us that the growth of white matter or myelination did not affect our segmentation over the large age span. The significant gender difference for the total brain volume is consistent with autopsy findings (39,40). When corrected for the brain size, no gender-specific differences in gray matter or white matter were noted.

As evident from the scatterplots, there is considerable normal variability of brain tissue volumes, which remains even after when the age, height, and weight are controlled (39,40). To overcome this variability, large data sets, as used, are needed to detect developmental changes. The method presented here, which can be applied accurately and reliably to segment large data sets of MR brain images, revealed developmentally significant segmentation results consistent with the earlier findings. Therefore, the method appears promising for studies of anatomic brain development and its relation to physiological and behavioral measures. It is being applied to a longitudinal brain morphological study.

APPENDIX A

Probabilistic Fusion Equation

For data from independent sensors s_1, s_2, . . . ,s_n, the probability of an event x, given data from all the sensors, $p(x|s_1, s_2, . . . ,s_n)$, is given by

$$p(x|s_1, s_2, . . . , s_n) =$$

$$k \cdot p_1(x|s_1) \cdot p_2(x|s_2) . . . p_n(x|s_n) \qquad (A1)$$

where k is constant depending on sensors and the event x and $p_r(x|s_r)$ denotes the probability of the event given the data from the rth sensor s_r.

Proof of Probabilistic Fusion Equation

Consider the inputs given by the data from two sensors s_1 and s_2. The probability of an event x given information from both sensors s_1 and s_2, $p(x|s_1, s_2)$, is as follows:

$$p(x|s_1, s_2) =$$

$$\frac{P(s_1, s_2|x) \cdot p(x)}{P(s_1, s_2)} \quad \text{from Bayes theorem} \qquad (A2)$$

$$= \frac{P_1(s_1|x) \cdot P_2(s_2|x) \cdot p(x)}{P(s_1, s_2)} \quad \text{since } s_1 \text{ and}$$

$$s_2 \text{ are independent} \qquad (A3)$$

$$= \frac{p_1(x|s_1) \cdot P_1(s_1)}{p_1(x)} \cdot \frac{p_2(x|s_2) \cdot P_2(s_2)}{p_2(x)}$$

$$\cdot \frac{p(x)}{P(s_1,s_2)} \quad \text{from Bayes theorem} \tag{A4}$$

$$= \frac{P_1(s_1) \cdot P_2(s_2) \cdot p(x)}{p_1(x) \cdot p_2(x) \cdot P(s_1,s_2)} \cdot p_1(x|s_1) \cdot p_2(x|s_2) \tag{A5}$$

$$= k \cdot p_1(x|s_1) \cdot p_2(x|s_2) \tag{A6}$$

where

$$k = \frac{P_1(s_1) \cdot P_2(s_2) \cdot p(x)}{p_1(x) \cdot p_2(x) \cdot P(s_1,s_2)}$$

is a constant for a given set of sensors and the event. $P(s_1,s_2)$ is prior joint probability between the sensors, and $P(s_1)$ and $P(s_2)$ are the prior probabilities of the sensors s_1 and s_2, respectively. $p_1(x)$ and $p_2(x)$ are the probabilities of event x seen by the sensors. Hence, the probabilistic fusion equation is true for two sensors, and similarly it can be proved for arbitrary n sensors. The Bayes theorem is stated and a more elaborate description of the theorem can be found in ref. 35.

Bayes Theorem: If a prior probability of a class ω is $P(\omega)$ and the conditional density of the event x given the class ω is $p(x|\omega)$, then the posterior probability $P(\omega|x)$ is given by

$$P(\omega|x) = \frac{p(x|\omega) \cdot P(\omega)}{p(x)} \tag{A7}$$

where $p(x)$ is the probability of the event.

REFERENCES

1. Miller AKH, Alston RL, Corsellis JAN. Variation with age in the volumes of grey and white matter in the cerebral hemispheres of man: measurements with an image analyzer. *Neuropathol Appl Neurobiol* 1980;6:119–32.

2. Thaler HT, Ferber PW, Rottenberg DA. A statistical method for determining the proportions of gray matter, white matter, and CSF using computed tomography. *Neuroradiology* 1978;16:133–5.

3. DeLeo JM, Schwartz M, Creasy H, Cutler N, Rapoport SI. Computer assisted categorization of brain computerized tomography pixels into cerebrospinal fluid, white matter, and gray matter. *Comput Biomed Res* 1985;18:79–88.

4. Andreasen NC. *Brain imaging: application in psychiatry.* Washington, D.C.: American Psychiatric Association, 1989: 67–121.

5. Haralick RM, Shapiro LG. Image segmentation techniques. *Comput Vis Graph Image Proc* 1985;29:100–32.

6. Bezdek JC, Hall LO, Clarke LP. Review of MRI image segmentation techniques using pattern recognition. *Med Phys* 1993;20:1033–48.

7. Kennedy DN, Filipek PA, Caviness VS Jr. Anatomic segmentation and volumetric calculations in nuclear magnetic resonance imaging. *IEEE Trans Med Imag* 1989;8:1–7.

8. Van Meter JW. Segmentation and restoration of magnetic resonance images using material mixture models. Ph.D. dissertation, Dartmouth College, 1993.

9. Cohen G, Andreasen NC, Alliger R, et al. Segmentation techniques for the classification of brain tissue using magnetic resonance imaging. *Psychiatry Res* 1992;45:33–51.

10. Jernigan TL, Press GA, Hesselink JR. Methods for measuring brain morphologic features on magnetic resonance images: validation and normal aging. *Arch Neurol* 1990;47:27–32.

11. Lim KO, Pfefferbaum A. Segmentation of MR brain images into cerebrospinal fluid spaces, white and gray matter. *J Comput Assist Tomogr* 1989;13:588–93.

12. Harris GJ, Barta PE, Peng LW, et al. MR volume segmentation of gray matter and white matter using manual thresholding: dependence on image brightness. *Am J Neuroradiol* 1994;15:225–30.

13. Raya SP. Low level segmentation of 3-D magnetic resonance brain images—a rule based system. *IEEE Trans Med Imag* 1990;9:327–37.

14. Gerig G, Martin J, Kikinis R, Kubler O, Shenton M, Jolesz FA. Automating segmentation of dual-echo MR head data. Proceedings of 12th International Conference on Information Processing in Medical Imaging, 1991:175–87.

15. O'Donnell M, Gore JC, Adams WJ. Toward an automated analysis system for nuclear magnetic resonance imaging. II. Initial segmentation algorithm. *Med Phys* 1986;13:293–7.

16. Brant ME, Bohan TP, Kramer LA, Fletcher JM. Estimation of CSF, white, and gray matter volumes in hydrocephalic children using fuzzy clustering of MR images. *Comput Med Imag Graph* 1994;18:25–34.

17. Vannier MW, Butterfield RL, Jordan D, Murphy WA, Levitt RG, Gado M. Multispectral analysis of magnetic resonance images. *Radiology* 1985;154:221–4.

18. Vannier MW, Pilgrim TK, Speidel TM, Neumann RL, Rickman DL, Schertz LD. Validation of magnetic resonance imaging multispectral tissue classification. *Comput Med Imag Graph* 1991;15:217–23.

19. Clarke LP, Velthuizen RP, Phuphanich S, Schellenberg JD, Arrington JA, Silbiger M. MRI: stability of three supervised segmentation techniques. *Magn Res Imag* 1993;11:95–106.

20. Amamoto DY, Kasturi R, Mamourian A. Tissue-type discrimination in magnetic resonance images. Proceedings of 10th International Conference on Pattern Recognition, 1990:603–7.

21. Liang Z. Tissue classification and segmentation of MR images. *IEEE Eng Med Biol* 1993;39:81–5.

22. Choi HS, Haynor DR, Kim Y. Partial volume tissue classification of multichannel magnetic resonance images. *IEEE Trans Med Imag* 1991;10:395–407.

23. Hyman TJ, Kurland RJ, Levy GC, Shoop JD. Characterization of normal brain tissue using seven calculated MRI parameters and a statistical analysis system. *Magn Res Med* 1989;11:22–34.

24. Mitiche A, Aggarwal JK. Multiple sensor integration/fusion through image processing: a review. *Opt Eng* 1986;25:380–6.

25. Luo RC, Kay MG. Multisensor integration and fusion in intelligent systems. *IEEE Trans Systems Man Cybernet* 1989;19:909–31.

26. Yakovlev PI, Lecours AR. The myelogenetic cycles of regional maturation of the brain. In: Minkowski A, ed. *Regional development of the brain in early life.* Oxford: Blackwell Scientific, 1963:3–69.

27. Holland BA, Hass DK, Norman D, Brandt-Zawadzki M, Newton TH. MRI of normal brain maturation. *Am J Neuroradiol* 1986;7:201–8.

28. Martin E, Kikinis R, Zuerrer M, et al. Developmental stages of human brain: an MR study. *J Comput Assist Tomogr* 1988;12:917–22.

29. Pfefferbaum A, Mathalon DH, Sullivan EV, Rawles JM, Zipursky RB, Lim KO. A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch Neurol* 1994;51:874–87.

30. Jernigan RL, Trauner DA, Hesselink JR, Tallal PA. Matu-

ration of human cerebrum observed in vivo during adolescence. *Brain* 1991;114:2037–49.

31. Giedd JN, Snell JW, Lange N, et al. Quantitative magnetic resonance imaging of human brain development. *Cereb Cortex* (in press).

32. *ANALYZE: reference manual, version 6.2.* Biomedical Imaging Resource, Mayo Foundation, Rochester, MN.

33. DeCarli C, Maisog J, Murphy DGM, Teichberg D, Rapoport SI, Horwitz B. Method for quantification of brain, ventricular, and subarachnoid CSF volume from MR images. *J Comput Assist Tomogr* 1992;16:274–84.

34. Ardekani BA, Braun M, Kanno I, Hutton BF. Automatic detection of intradural spaces in MR images. *J Comput Assist Tomogr* 1994;18:963–9.

35. Duda RO, Hart PE. *Pattern classification and scene analysis.* New York: Wiley-Interscience, 1973.

36. Kohn MI, Tanna NK, Herman GT, et al. Analysis of brain and cerebrospinal fluid volumes with MR imaging, part 1: methods, reliability and validation. *Radiology* 1991;178:115–22.

37. Bulte JW, Douglas T, Mann S, et al. Magnetoferritin: characterization of a novel superparamagnetic MR contrast agent. *J Magn Res Imag* 1994;4:497–505.

38. Caviness VS, Filipek PA, Kennedy DN. Magnetic resonance technology in human brain science: blueprint for a program based upon morphometry. *Brain Devel* 1989;11:1–13.

39. Dekaban AS, Sadowsky D. Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Ann Neurol* 1978;4:345–56.

40. Ho K, Roessmann U, Straumfjord JV, Monroe G. Analysis of brain weight: adult brain weight in relation to sex, race and age. *Arch Pathol Lab Med* 1980;104:635–49.

41. Jacobson M. *Developmental neurobiology.* 3rd ed. New York: Plenum Press, 1991.